

High precision series solution of differential equations: Ordinary and regular singular point of second order ODEs.

Amna Noreen^a, Kåre Olaussen^a

^aInstitutt for fysikk, NTNU

Abstract

A subroutine for very-high-precision numerical solution of a class of ordinary differential equations is provided. For given evaluation point and equation parameters the memory requirement scales linearly with precision P , and the number of algebraic operations scales roughly linearly with P when P becomes sufficiently large. We discuss results from extensive tests of the code, and how one for a given evaluation point and equation parameters may estimate precision loss and computing time in advance.

Keywords:

Second order ODEs, Regular singular points, Ordinary points, Frobenius method.

PROGRAM SUMMARY

Manuscript Title: High precision series solution of differential equations: Ordinary and regular singular point of second order ODEs.

Authors: Amna Noreen, Kåre Olaussen

Program Title: seriesSolveOde1

Journal Reference:

Catalogue identifier:

Licensing provisions: none

Programming language: C++

Computer: PC's or higher performance computers

Operating system: Linux and MacOS

RAM: Few to many megabytes (problem dependent)

Number of processors used: 1

Keywords: Second order ODEs, Regular singular points, Ordinary points, Frobenius method.

Classification: 2.7 Wave functions and integrals, 4.3 Differential equations.

External routines/libraries: CLN – Class Library for Numbers [1] built with the GNU MP library [2], and GSL – GNU Scientific Library [3] (only for time measurements).

Subprograms used: The code of the main algorithm is in the file `seriesSolveOde1.cc`, which `#include` the file `checkForBreakOde1.cc`. These routines, and programs using them, must `#include` the file `seriesSolveOde1.cc`.

Nature of problem: The differential equation

$$-s^2 \left(\frac{d^2}{dz^2} + \frac{1 - \nu_+ - \nu_-}{z} \frac{d}{dz} + \frac{\nu_+ \nu_-}{z^2} \right) \psi(z) + \frac{1}{z} \sum_{n=0}^N \nu_n z^n \psi(z) = 0, \quad (1)$$

is solved numerically to very high precision. The evaluation point z and some or all of the equation parameters may be complex numbers; some or all of them may be represented exactly in terms of rational numbers.

Solution method:

The solution $\psi(z)$, and optionally $\psi'(z)$, is evaluated at the point z by

executing the recursion

$$A_{m+1}(z) = \frac{s^{-2}}{(m+1+\nu_+)(m+1+\nu_--\nu_-)} \sum_{n=0}^N V_n(z) A_{m-n}(z), \quad (2)$$

$$\psi^{(m+1)}(z) = \psi^{(m)}(z) + A_{m+1}(z), \quad (3)$$

to sufficiently large m . Here ν is either ν_+ or ν_- , and $V_n(z) = \nu_n z^{n+1}$. The recursion is initialized by

$$A_{-n}(z) = \delta_{n0} z^{\nu}, \quad \text{for } n = 0, 1, \dots, N \quad (4)$$

$$\psi^{(0)}(z) = A_0(z). \quad (5)$$

Restrictions: No solution is computed if $z = 0$, or $s = 0$, or if $\nu = \nu_-$ (assuming $\text{Re } \nu_+ \geq \text{Re } \nu_-$) with $\nu_+ - \nu_-$ an integer, except when $\nu_+ - \nu_- = 1$ and $\nu_0 = 0$ (i.e. when z is an ordinary point for $z^{-\nu_-} \psi(z)$).

Running time: On an few years old Linux PC, evaluating the ground state wavefunction of the anharmonic oscillator (with the eigenvalue known in advance), cf. equation (6), at $y = \sqrt{10}$ to $P = 200$ decimal digits accuracy takes about 2 milliseconds, to $P = 100\,000$ decimal digits accuracy takes about 40 minutes.

References

- [1] B. Haible and R.B. Kreckel, *CLN – Class Library for Numbers*, <http://www.ginac.de/CLN/>
- [2] T. Granlund and collaborators, *GMP – The GNU Multiple Precision Arithmetic Library*, <http://gmplib.org/>
- [3] M. Galassi *et al*, *GNU Scientific Library Reference Manual* (3rd Ed.), ISBN 0954612078., <http://www.gnu.org/software/gsl/>

1. Introduction

Modelling and analysis of many problems in science and engineering involves the solution of ordinary differential equations, sometimes in the domain of complex numbers. For practical use such solutions must usually be computed numerically at some stage. In some cases it may be necessary, useful, or

Email addresses: Amna.Noreen@ntnu.no (Amna Noreen),
Kare.Olaussen@ntnu.no (Kåre Olaussen)

interesting to do this to much higher precision than provided by standard equation solvers (or routines for evaluating standard functions).

Some examples of cases where numerical calculations have been used to inspire or check analytic conjectures and proofs are the works by Bender and Wu [1] and Zinn-Justin and Jentschura [2]. With access to very accurate numerical results the opportunities for such explorations increases. Access to essentially exact results are also useful for analyzing the behaviour of approximation schemes, as in the work by Bender *et al* [3] and more recently by Mushtaq *et al* [4].

We have implemented and investigated the algorithm (2, 3) for solving equation (1) to very high precision, and believe the C++ function `seriesSolveOde1` may be of use or interest to others. An early version of this code has been used to solve eigenvalue problems like the anharmonic oscillator and the double well potential,

$$\left[-\frac{d^2}{dy^2} + y^4\right]\psi(y) = \varepsilon_n \psi(y), \quad (6)$$

$$\left[-s^2 \frac{d^2}{dy^2} + (1 - y^2)^2\right]\psi(y) = \varepsilon_{n\sigma} \psi(y), \quad (7)$$

to very high precision. In reference [5] the ground state eigenenergy ε_0 of (6) was found to 1^+ million decimal digits precision, the excited state $\varepsilon_{50\,000}$ was solved to $50\,000^+$ decimals, and the lowest even, ε_{0+} , and odd parity, ε_{0-} , eigenvalues of (7), with $s = 1/50\,000$, was found to $30\,000^+$ decimals. Equations (6, 7) are transformed to the form (1) by introducing $z = y^2$, leading to $\nu_- = 0$ and $\nu_+ = \frac{1}{2}$ and $\nu_2 = \frac{1}{4}$. This further gives $\nu_0 = -\frac{1}{4}\varepsilon_n$ for equation (6), and $\nu_0 = \frac{1}{4}(1 - \varepsilon_{n\sigma})$, $\nu_1 = -\frac{1}{2}$ for equation (7). I.e., the eigenvalue parameter enters in the coefficient $A_0(z)$ of equation (8).

The eigenvalue condition for these problems is that the wave function should vanish as $y \rightarrow \pm\infty$, a condition which cannot be imposed numerically. However, an asymptotic analysis of the behaviour of the wavefunction as $y \rightarrow \pm\infty$ allows us to replace it with an equivalent Robin boundary condition at finite y . The latter cannot be computed exactly, but to sufficient accuracy for any desired precision. In fact, if we make y large enough it suffices to use a Diriclet boundary condition.

In reference [6] it was demonstrated that the wavefunction normalization integrals can be computed to comparable precision, again using an early version of our code.

In the rest of this paper we provide examples of how this code can be used, and some analysis of its behaviour. We do not focus on specific areas of applications, but would like to mention that very-high-precision computations of Green functions and functional determinants are possible applications. The code can evaluate $\psi(z)$ for complex values of z , and allow for complex parameters in the differential equation.

2. Basic use

The function `seriesSolveOde1` is declared as

Function declaration

```
bool seriesSolveOde1(OdeResults& results, const
    OdeParams& params)
```

The function returns `true` if the calculation completed normally and `false` if the calculation was aborted. Function arguments and options are collected in a structure `OdeParams`, with the function value $\psi(z)$, optionally $\psi'(z)$, and various diagnostic results returned in a structure `OdeResults`. The definitions of these structures are listed in the Appendix at the end of this paper. A code snippet illustrating the use of `seriesSolveOde1` is

Example use

```
// Better start from default parameters
OdeParams params = defaultOdeParams;
// Change parameters as needed
params.prec = float_format(500);
params.z = complex(27/2, 43/7);
params.dAlso = true;
OdeResults results;
// NB! Must allocate space for  $\psi(z)$  a  $\psi'(z)$ 
cl_N fu[2];
results.fu = fu;
if(seriesSolveOde1(results, params)){
    cout << results.fu[0] << endl;
    cout << results.fu[1] << endl;
}
```

3. Computational accuracy

Our solution is found by a *brute force* summation,

$$\psi(z) = \sum_{m=0}^{\infty} a_m z^{\nu+m} \approx \sum_{m=0}^M a_m z^{\nu+m} \equiv \sum_{m=0}^M A_m(z). \quad (8)$$

Since equation (1) has no singularities in the finite z -plane, except perhaps $z = 0$, the sum is guaranteed to have an infinite radius of convergence. However, intermediate terms in the sum may be very large although the final result is small; hence there may be huge cancellations, leading to significant loss of accuracy when z is large.

The computations are performed with high-precision floating point numbers, with precision regulated by the parameter `params.prec` as shown in the code snippet above. The value given is the intended precision P in decimal digits, but since memory for floating point numbers is allocated internally (in CLN at high precision) in chunks of 64 bits the actual precision is usually somewhat higher — increasing in steps of about $19 \approx 64 \cdot \log 2 / \log 10$ decimal digits.

Denote the actual precision used in computation by M . I.e., each non-zero floating point number x is represented as $(-1)^s 2^m f$, with the mantissa f ($\frac{1}{2} \leq f < 1$) given to a precision of M bits. This means that the potential roundoff error in x is 2^{m-M-1} . I.e, if the largest term $A_m(z)$ in the sum (8) has the representation $(-1)^s 2^{\bar{A}} f$ it may contribute a roundoff error $2^{\bar{A}-M-1}$ to ψ . Due to the recursion relation (2) roundoff errors may be

further amplified (or partially cancelled), but it is a reasonable hypothesis to use \bar{A} for an estimate of the evaluation error.

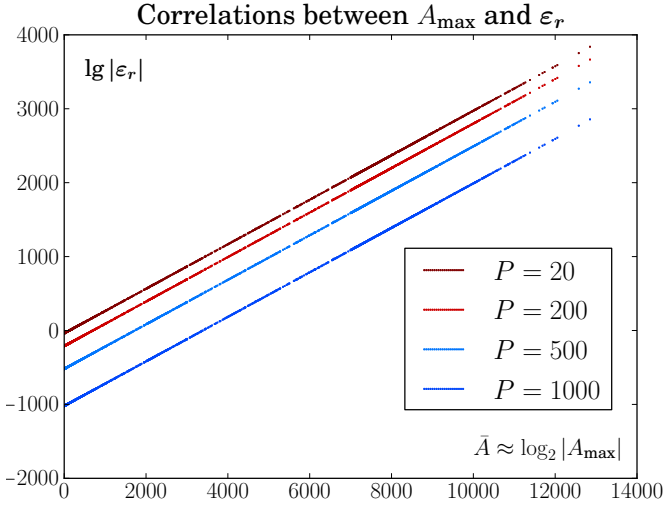


Figure 1: The real evaluation error, $\varepsilon_r \equiv |\psi_P(z) - \psi_E(z)|$, caused by roundoff is strongly correlated with the largest term A_{\max} in the series expansion. Here $\psi_P(z)$ is the value obtained when evaluating the series to the intended precision of P decimal digits, while $\psi_E(z)$ is the “exact” value (here actually the value obtained for $P = 1040$).

We have tested this hypothesis by running a large number (200 000) of evaluations with randomly chosen parameters, and investigated the correlation between \bar{A} and the real evaluation error $\varepsilon_r \equiv |\psi_P(z) - \psi_E(z)|$ for various precisions P . Here $\psi_E(z)$, representing the exact value, is found by doing the computation with a precision P_E reasonably larger than all the others. As can be seen in figure 1 the correlation between \bar{A} and ε_r is good compared to the accuracies in question.

For diagnostic purposes the value of \bar{A} is returned by `seriesSolveOde1` in the variable `results.maxAExponent`. The corresponding value for $\psi'(z)$ is returned in the variable `results.maxAdExponent` (the values of m where the maxima occur are also returned). Based on these values and the actual precision M the estimated errors in decimal digits are returned in the variables `results.lgErrorF` and `results.lgErrorFd`. Their exact values are $(\bar{A} - M) \log 2 / \log 10 + G$ for ψ , and $(\bar{A}' - M) \log 2 / \log 10 + G'$ for ψ' , where the numbers $G = 4.30$ and $G' = 3.02$ are empirically chosen “guard digits” to avoid underestimating the error (too often).

These estimates are accurate to a handful of decimal digits as shown by the $\Delta = \lg |\varepsilon_r| - \lg |\varepsilon_e|$ histogram in figure 2. We have found such histograms to be independent of computational precision P , and (with the chosen value of $G - G'$) also the same for $\psi(z)$ and $\psi'(z)$. The histogram is taken over 200 000 evaluations with N chosen randomly between 1 and 4, the real and imaginary parts of s randomly from the set $\{-1, -\frac{1}{3}, \frac{1}{3}, 1\}$, the real and imaginary parts of v_{\pm} randomly between -10 and 10 , the real and imaginary parts of v_n randomly between -5 and

5, and the real and imaginary parts of z randomly between -20 and 20 .

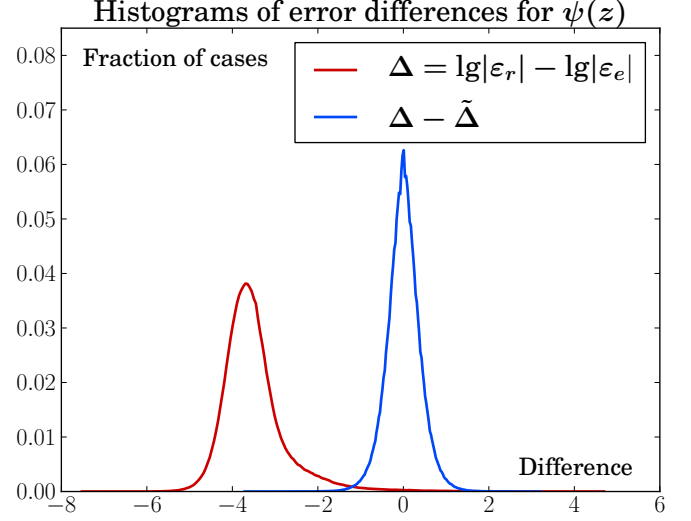


Figure 2: The curve for $\Delta = \lg |\varepsilon_r| - \lg |\varepsilon_e|$ shows a histogram of differences between the real evaluation error ε_r , and an estimate based on the largest term A_{\max} in the sum (8). The histogram does not depend on computational precision P , but individual differences (with fixed z and equation parameters) varies with P in a fluctuating manner. The curve for $\Delta - \tilde{\Delta}$ shows how differences computed at precision $P = 20$ ($\tilde{\Delta}$) correlates with those computed at $P = 200, 500, 1000$.

The difference Δ is caused by an essentially unpredictable roundoff error, amplified by a recursion relation which depends on z and parameters of the differential equation. As can be seen from figure 2 the ratio between the real and estimated errors varies between almost 10^5 and 10^{-8} . Although this variation is large it is still less than the discrete steps by which the actual precision is increased.

We have investigated how the difference Δ computed at different precisions (but for the same evaluation point z and equation parameters) are correlated. This is shown in the $\Delta - \tilde{\Delta}$ histogram in figure 2, where $\tilde{\Delta}$ refer to a computation with intended precision $P = 20$, and Δ to computations with $P = 200, 500$, and 1000 (where each P gives the same looking histogram). As can be seen the correlations are stronger than for a single Δ , but there are still wide tails.

In conclusion, the largest term A_{\max} in the series (8), or the associated integer \bar{A} , provides a good estimate of the evaluation error, but the real error may still differ by several orders of magnitude. As somewhat better empirical estimate can be obtained by first computing the real error at low P (where it is computationally inexpensive) and assuming

$$\lg \varepsilon_E = \lg \tilde{\varepsilon}_r - (M - \tilde{M}) \log 2 / \log 10 + 2. \quad (9)$$

Here ε_r is the real error at \tilde{M} bits of actual precision, with ε_E the estimated error at M bits of actual precision.

4. Comparison with exactly known Wronski determinant

In the previous section we assumed that `seriesSolveOde1` would compute accurate results at large intended precision P , but this was not really verified. One check is to compare its results with the exactly known Wronski determinant,

$$W(z) = \psi_{v_+}(z) \psi'_{v_-}(z) - \psi_{v_-}(z) \psi'_{v_+}(z) = (v_- - v_+) z^{v_+ + v_- - 1}. \quad (10)$$

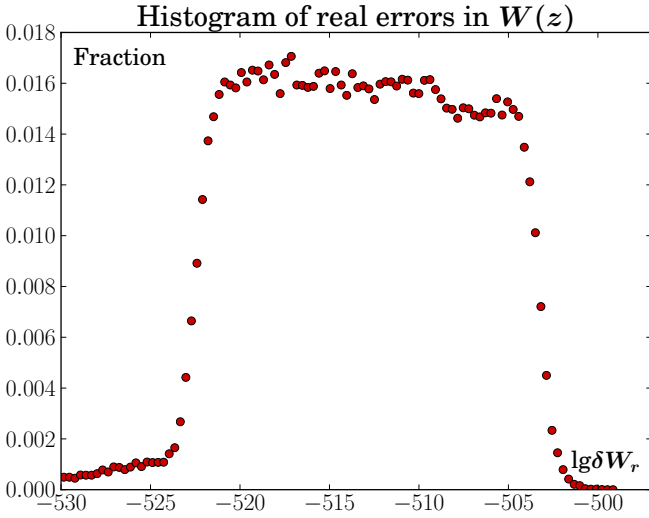


Figure 3: The real errors $\delta W_r \equiv |W^{\text{exact}}(z) - W^{\text{num}}(z)|$ when aiming to compute the Wronski determinant $W(z)$ to an absolute accuracy of 10^{-500} . The histogram is taken over 150 000+ random evaluation points z and equation parameters.

The numerically computed determinant is estimated to have an error of magnitude

$$\delta W_e = \max \left\{ |\psi_{v_+}| \delta \psi'_{v_-}, |\psi_{v_-}| \delta \psi'_{v_+}, |\psi'_{v_-}| \delta \psi_{v_+}, |\psi'_{v_+}| \delta \psi_{v_-} \right\}, \quad (11)$$

where f.i. $\delta \psi'_{v_-}$ is the estimated magnitude of error in ψ'_{v_-} (all quantities evaluated at z). An estimate of the loss of precision can be made by a calculation at low(er) P , and used to choose the appropriate value of `params.prec` for a desired final precision in $W(z)$. Figure 3 shows a histogram of how this works, tested on a large number of random evaluation points and equation parameters. In most cases the real precision is reasonably close (always better) than the one aimed for, but sometimes it turns out to be much better. This may occur when the low precision calculation overestimates the magnitude of $\psi(z)$ or $\psi'(z)$. However, as shown in figure 4 the real error δW_r in the numerically computed determinant is always reasonably close to the final estimate δW_e based on equation (11) with all quantities computed at high precision.

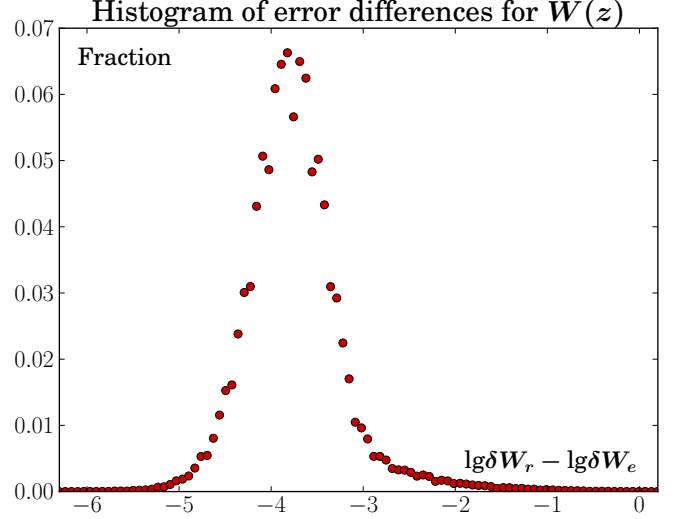


Figure 4: Histogram of differences between the real error δW_r in the numerically computed Wronski determinant, and the estimated error δW_e based on (11).

The conclusion is that the calculated Wronski determinants are correct within the expected accuracies, at least for computations at sufficiently high precision (how high may depend on the evaluation point z and equation parameters).

5. A priori accuracy estimates

Although `seriesSolveOde1` monitors the largest term in the sum (8) to estimate the accuracy of the computed results, it is desirable to predict the behaviour of the sum in advance. One way to do so is by analysing the behaviour of the recursion relation (2). Quite detailed and interesting results can be found in simple cases, but the analysis becomes unmanageable in general. We instead make the hypothesis that the terms in the sum (8) for large $z = x e^{i\varphi}$ is strongly peaked (in absolute value) around some $m = \bar{m}$, and that there exist values of φ for which there are little cancellation between the large terms. I.e., we assume that

$$\max_{\varphi} |\psi(x e^{i\varphi})| \approx |a_{\bar{m}}| x^{\nu + \bar{m}}, \quad (12)$$

for positive x . We may use the WKB-approximation to estimate the left hand side. For analytic treatment we first neglect the slowly varying algebraic prefactor of the WKB-approximation, and a similar correction to the relation (12). Such corrections can be included in a numerical implementations.

Define, for positive u ,

$$S(u) = \max_{\varphi} \log |\psi(e^{u+i\varphi})| \approx \max_{\varphi} \operatorname{Re} \left(\int_0^{e^{u+i\varphi}} Q(t) dt \right), \quad (13)$$

where $Q(t)$ is found from the differential equation (1). We then have the relation

$$S(u) = \log(|a_{\bar{m}}|) + (\nu + \bar{m}) u, \quad (14)$$

$$u = -\frac{d}{dm} \log(|a_m|) \Big|_{m=\bar{m}}. \quad (15)$$

The last equation follows from the maximum condition. We recognize (14, 15) as a Legendre transform [[8],[9], [10]]. By inverting this transform we find

$$\bar{m} = \frac{d}{du} S(u), \quad \log(|a_{\bar{m}}|) = S(u) - (\nu + \bar{m})u, \quad (16)$$

which provides an *a priori* order-of-magnitude estimate of the coefficients a_m , and hence of (i) the accuracy loss due to numerical roundoff, and (ii) the number of terms \mathcal{M} required in (8) for a desired final precision.

5.1. Example 1: Anharmonic oscillators

Consider the equation

$$-\frac{\partial^2}{\partial y^2} \Psi(y) + (y^2 + c^2)^2 \Psi(y) = 0. \quad (17)$$

For large y the typical solution behaves like

$$\Psi(y) \sim e^{\frac{1}{3}y^3 + c^2 y}, \quad (18)$$

neglecting the slowly varying prefactor. Equation (17) can be transformed to the form (1) by introducing $x = y^2$, $\Psi(y) = \psi(x)$. Hence, with $x = y^2 = e^u$

$$S(u) = \frac{1}{3} \left(e^{\frac{3}{2}u} + 3c^2 e^{\frac{1}{2}u} \right),$$

which gives

$$\bar{m} = \frac{1}{2} \left(e^{\frac{3}{2}u} + c^2 e^{\frac{1}{2}u} \right), \quad (19)$$

$$\log(|a_{\bar{m}}|) = \left(\frac{1}{3} - \frac{1}{2}u \right) e^{\frac{3}{2}u} + c^2 \left(1 - \frac{1}{2}u \right) e^{\frac{1}{2}u}. \quad (20)$$

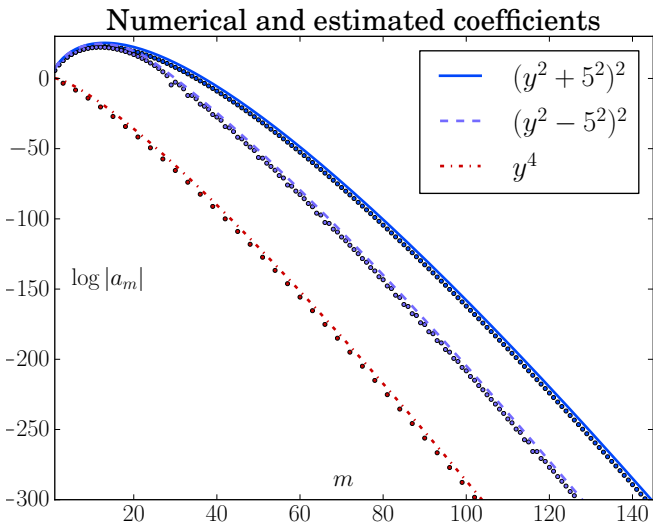


Figure 5: Comparison of numerical coefficients a_m (points) with estimates (full-drawn lines) based on (19, 20) and (28, 29). The estimates of $\log|a_m|$ are accurate up to corrections which depend logarithmically on m .

Note that (19, 20) give, in parametric form, an estimate of *all* coefficients $|a_m|$, not only those corresponding to a maximum value. For $c = 0$ an explicit representation is easily found to be

$$\log|a_m| = \frac{2}{3}m(1 - \log 2m). \quad (21)$$

This is plotted as the lower curve in figure 5. It fits satisfactory with the high-precision coefficients generated numerically, but there remains a correction which depends logarithmically on m . For nonzero c the parametric representation provides equally good results, as shown by the upper curve in figure 5.

The conclusion of this example is that we expect the largest term of the power series to be

$$\max_m |A_m(x)| \sim e^{\frac{1}{3}(x^{3/2} + 3c^2 x^{1/2})}, \quad (22)$$

neglecting a slowly varying prefactor. Further, the maximum should occur at

$$m \approx \frac{1}{2} \left(x^{3/2} + c^2 x^{1/2} \right). \quad (23)$$

Finally, estimates like equation (21) for the coefficients a_m may be used to predict how many terms \mathcal{M} we must sum to evaluate $\psi(x)$ to a given precision P , based on the stopping criterium

$$|a_M| x^M \leq 10^{-P}. \quad (24)$$

As can be seen in figure 6 the agreement with the actual number of terms used by `seriesSolve0de1` is good, for case of equation (17) with $c = 0$, in particular for high precision P . But note that a logarithmic scale makes it easier for a comparison to look good.

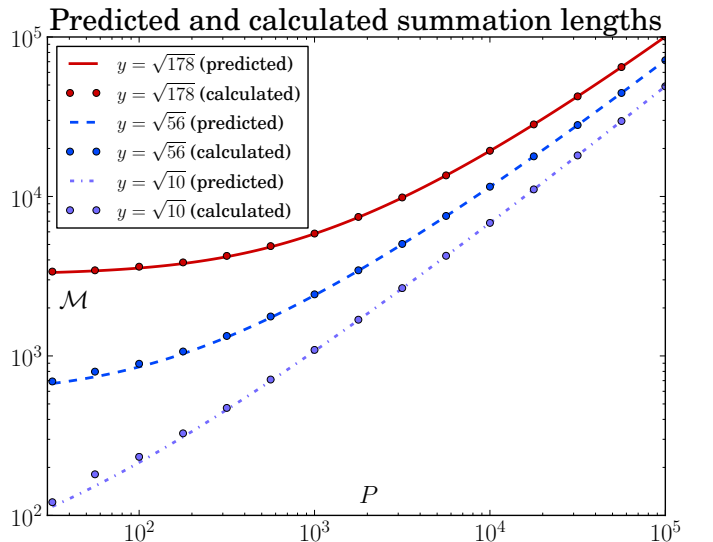


Figure 6: This figure compares the *a priori* prediction, based on equation (21), of the number of terms \mathcal{M} which must be summed in order to evaluate $\Psi(y)$ for $c = 0$ to a desired precision P with the actual number of terms computed by `seriesSolve0de1`.

5.2. Example 2: Double well oscillators

Next consider the equation

$$-\frac{\partial^2}{\partial y^2}\Psi(y) + (y^2 - c^2)^2 \Psi(y) = 0. \quad (25)$$

For large y the typical solution behaves like

$$\Psi(y) \sim e^{\frac{1}{3}y^3 - c^2 y}, \quad (26)$$

neglecting the slowly varying prefactor. Equation (25) can be transformed to the form (1) by introducing $x = y^2$, $\Psi(y) = \psi(x)$. Hence, with $x = y^2 = e^u$

$$S(u) = \max_{\varphi} \frac{1}{3} \operatorname{Re} \left(e^{\frac{3}{2}(u+i\varphi)} - 3c^2 e^{\frac{1}{2}(u+i\varphi)} \right).$$

The maximum occurs for $\cos \frac{1}{2}\varphi = -\frac{1}{2} \left(1 + c^2 e^{-u} \right)^{1/2}$ when $e^u \geq \frac{1}{3}c^2$, and for $\cos \frac{1}{2}\varphi = -1$ otherwise. This gives

$$S(u) = \begin{cases} c^2 e^{u/2} - \frac{1}{3} e^{3u/2} & \text{for } e^u \leq \frac{1}{3}c^2, \\ \frac{1}{3} (e^u + c^2)^{3/2} & \text{for } e^u \geq \frac{1}{3}c^2. \end{cases} \quad (27)$$

This implies that

$$\bar{m} = \begin{cases} \frac{1}{2} e^{u/2} (c^2 - e^u) & \text{for } e^u \leq \frac{1}{3}c^2, \\ \frac{1}{2} e^u (e^u + c^2)^{1/2} & \text{for } e^u \geq \frac{1}{3}c^2, \end{cases} \quad (28)$$

$$\log(|a_{\bar{m}}|) = \begin{cases} \left(1 - \frac{1}{2}u \right) c^2 e^{u/2} - \left(\frac{1}{3} - \frac{1}{2}u \right) e^{3u/2} & \text{for } e^u \leq \frac{1}{3}c^2, \\ \left[\left(\frac{1}{3} - \frac{1}{2}u \right) e^u + \frac{1}{3}c^2 \right] (e^u + c^2)^{1/2} & \text{for } e^u \geq \frac{1}{3}c^2. \end{cases} \quad (29)$$

This representation compares fairly well with the numerically generated coefficients, as shown by the middle curve in figure 5. However, in this case the coefficients a_m have a local oscillating behaviour. The representation (28, 29) should be interpreted as the local amplitude of this oscillation.

The conclusion of this example is that we expect the largest term of the power series to be term of the series to be

$$\max_m |A_m(x)| \sim e^{\frac{1}{3}(x+c^2)^{3/2}}, \quad (30)$$

neglecting the slowly varying prefactor. Further, the maximum should occur at

$$m \approx \frac{1}{2}x(x+c^2)^{1/2} \approx \frac{1}{2}x^{3/2} + \frac{1}{4}c^2 x^{1/2}. \quad (31)$$

5.3. Logarithmic corrections

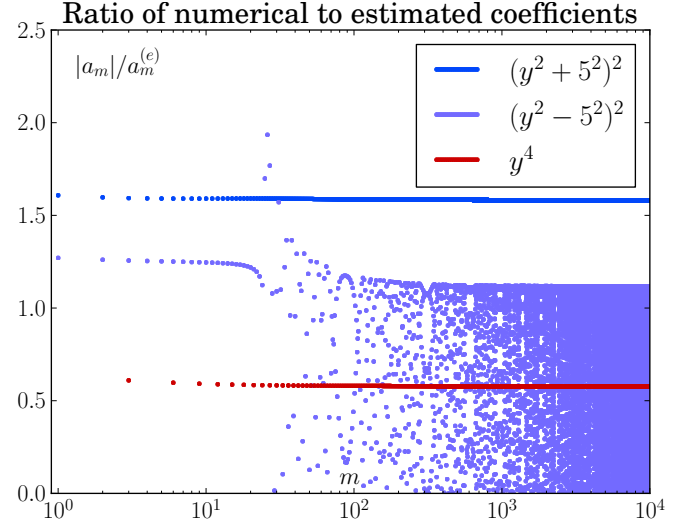


Figure 7: The ratio between the numerically generated $|a_m|$ and their estimated values $a_m^{(e)}$ based on equation (16) with $S(u)$ replaced by $S_{\text{eff}}(u)$. The important feature here is that the ratios are essentially constant, not which constant, due to an overall normalization constant which we have not attempted to estimate here.

The agreement between the *a priori* estimated magnitude $a^{(e)}$ and the numerically generated values $|a_m|$ looks quite good in figure 5. This is partly due to the logarithmic scale; on closer examination the coefficients are seen to differ by many orders of magnitude. The agreement can be significantly improved by (i) taking into account the prefactor $Q(u)^{-1/2}$ in the WKB-expression (13) for the left hand side of (12), and (ii) changing the right hand side of (12) as

$$|a_{\bar{m}}| x^{\nu+\bar{m}} \equiv e^{s(\bar{m})+(\nu+\bar{m})u} \rightarrow \sqrt{2\pi S''(u)} e^{s(\bar{m})+(\nu+\bar{m})u}.$$

The latter replacement takes into account that the main contribution to the sum over m comes from a range of values around \bar{m} , approximates this contribution by a gaussian integral, and uses the fact that $s''(\bar{m})^{-1} = -S''(u)$. These two improvements amounts to a change $S(u) \rightarrow S_{\text{eff}}(u)$.

Implementing these changes for the $c = 0$ case of (17), taking into account that only coefficients $a_{3n} \neq 0$, the estimate (21) can be improved to

$$\log a_m^{(e)} = \frac{2}{3} \left(m + \frac{5}{4} \right) \left[1 - \log(2m + \frac{5}{2}) \right] - \frac{1}{2} \log \frac{\pi}{6}. \quad (32)$$

Similar, algebraically more complicated, improvements can be made for $c > 0$. As shown in figure 7 the improved estimates compare very well with the numerically generated coefficients, but in case of locally oscillating a_m the (smooth) estimate $a_m^{(e)}$ should be interpreted as the local oscillation amplitude (as illustrated by the $(y^2 - 5^2)^2$ -case in figure 7).

6. Concluding remarks

For equations without singular points in the finite plane our code can be used for expansion around any point ζ_0 in the complex plane. To shift from one expansion point to another one just has to rewrite the parameters v_n , and let z denote the distance from ζ_0 . This allows for analytic continuation of the solution, which becomes quite easy since the full solution is determined by just the two parameters $\psi(\zeta_0)$, $\psi'(\zeta_0)$ (in addition to the differential equation).

The strategy of using a sequence of series expansions, each with a small parameter z , has been used by Haftel *et al* [11]. The advantage is that each summation requires fewer terms in the series, and may lead to less loss of precision caused by roundoff errors. The cost is of course that one has to do several sums, and one may also lose symmetries like the $y \rightarrow -y$ symmetry in equations (6, 7). The latter leads to more algebraic operations per recursion step.

The optimal strategy may depend on the problem. If we are only solving equation (6) for the ground state eigenvalue ε_0 , this is basically determined by the condition that the asymptotic behaviour of the solution switches very rapidly between $e^{y^3/3}$ and $-e^{y^3/3}$. This behaviour is not affected much by roundoff errors. Consider the question is whether it is faster to evaluate $e^{y^3/3}$ by a single series expansion, or by k expansions with $y_k = y/k$. By combining equations (21, 24) one finds that each sum requires about M_k terms for a given precision P , where M_k satisfies the equation

$$\frac{2}{3} M_k (1 - \log 2M_k) + 2M_k \log(y/k) \approx -P \log 10, \quad (33)$$

which is best solved numerically. Consider f.i. the case of $y = \sqrt{178}$ and $P = 10^5$. As can be seen from figure 6 about $M \equiv M_1 = 10^5$ terms have to be summed to obtain the desired precision. With $k = 2$ only about $M_2 = 67\,500$ terms has to be summed, but since this has to be done twice the total effort becomes larger. The situation is similar for other values of y and P .

In other cases, like highly excited states of (6) or all states of (7), there is a loss of precision due to roundoff. This changes how the number of terms M and the actual precision M vary with y . The latter is most important since the time per multiplication increases somewhat faster than quadratic with M . In such cases a sequence of analytically continued evaluations are clearly advantageous; optimization of the number and size of steps requires some prior knowledge of the coefficients a_n and the behaviour of the solution [13]. If one needs to evaluate the solution at a sequence of points, as when calculating the normalization integral [6], analytic continuation would also be preferable.

The routine `seriesSolveOde1` does not allow for analytic continuation in the presence of a regular singular point, since the transformed equation belong to a different class. We have developed and are testing code for a more general class of equations (as hinted by our naming scheme), which we intend to submit real soon. This code allow for translations (or more generally Möbius transformations) to a new expansion point ζ_0 . It can f.i. be used to solve Mathieu and Mathieu-like equations.

We have made extensive tests of the submitted code, which appears to be robust and perform according to theoretical expectations. As illustrated, surprisingly accurate *a priori* estimates of the series to be summed can be made by using the WKB approximation in combination with Legendre transforms. This is useful for estimating precision and time requirements in advance. In the general case the WKB integral and Legendre transformation must be computed numerically. We have developed and tested code for this purpose [12, 13].

Acknowledgement

We thank A. Mushtaq and I. Øverbø for useful discussions. This work was supported in part by the Higher Education Commission of Pakistan (HEC).

References

- [1] C.M. Bender and T.T. Wu, *Anharmonic oscillator*, Physical Review **184**, 1231–1260 (1971)
- [2] J. Zinn-Justin and U.D. Jentschura, *Multi-Instantons and Exact Results II: Specific Cases, Higher-Order Effects, and Numerical Calculations*, Annals Phys. **313** 269–325 (2004)
- [3] C.M. Bender, K. Olaussen and P.S. Wang, *Numerological analysis of the WKB approximation in large order*, Physical Review **D16**, 1740–1748 (1977)
- [4] A. Mushtaq, A. Kværnø, and K. Olaussen, *Systematic improvement of splitting methods for the Hamilton equations*, contribution to World Congress of Engineers 2012 (London, UK., 4–6 July, 2012)
- [5] A. Mushtaq, A. Noreen, K. Olaussen, and I. Øverbø, *Very-high-precision solutions of a class of Schrödinger type equations*, Computer Physics Communications **189**, 1810–1813 (2011)
- [6] A. Noreen and K. Olaussen, *Very-high-precision normalized eigenfunctions for a class of Schrödinger type equations*, World Academy of Science, Engineering and Technology: An International Journal of Science, Engineering and Technology **76**, 831–836 (2011)
- [7] F.G. Frobenius, *Über die Integration der linearen Differentialgleichungen durch Reihen*, Journal für die reine und angewandte Mathematik, **76**, 214 (1873)
- [8] R.T. Rockafellar, *Convex Analysis* paperback ed., Princeton University Press (1996)
- [9] K. Huang, *Statistical Mechanics* 2nd ed., John Wiley & Sons (1987)
- [10] R. K. P. Zia, Edward F. Redish, and Susan R. McKay, *Making Sense of the Legendre Transform*, [arXiv.org/0806.1147](https://arxiv.org/abs/0806.1147) (2008)
- [11] M. Haftel, R. Krivec, and V.B. Mandelzweig, *Power Series Solution of Coupled Differential Equations in One Variable*, Journal of Computational Physics **123**, 149–161 (1996)
- [12] A. Noreen and K. Olaussen, *Estimating coefficients of Frobenius series by Legendre transform and WKB approximation*, contribution to World Congress of Engineers 2012 (London, UK., 4–6 July, 2012)
- [13] A. Noreen and K. Olaussen, *Estimating coefficients for series solutions of differential equations: Ordinary and regular singular points of second order ODEs.*, separate submission to Computer Physics Communications.
- [14] A. Mushtaq, A. Noreen, K. Olaussen, and I. Øverbø, In preparation.

Appendix

Definition of the OdeParams structure

```

struct OdeParams {
    // Parameters defining the ODE
    cl_N z;           // Evaluation point
    cl_N nuP;         //  $\nu_+$ 
    cl_N nuM;         //  $\nu_-$ 
    cl_N s;
    int orderN;       // Order of polynomial potential
    cl_N* v;          // Pointer to array v[orderN+1]
    // Options regulating the computation
    float_format_t prec; // Wanted computational precision
    bool nuPlus;       // If true, compute  $\nu = \nu_+$  solution
    bool dAlso;        // If true, compute derivative also
    cl_I emmMax;       // Stop summation when emm=emmMax
    cl_I emmTooLarge;  // Stop if emm >= emmTooLarge
    // Options to print double precision coeff info to stdout
    bool writeParams;  // Write to (temporary) file?
    bool deleteParams; // Delete above file at normal exit?
    bool printLogAbsA;  // Option to print all  $\log(\text{abs}(A_m))$ 
    bool printLogReA;   // Option to print  $\log(\text{abs}(\text{Re}(A_m)))$ 
    bool printLogImA;   // Option to print  $\log(\text{abs}(\text{Im}(A_m)))$ 
    bool printArgA;     // Option to print  $\arg(A_m)$ 
    bool printReA;      // Option to print  $\text{Re}(A_m)$ 
    bool printImA;      // Option to print  $\text{Im}(A_m)$ 
    bool printLogAbsAd; // Print all  $\log(\text{abs}((\nu + m)A_m/z))$ 
    bool printLogReAd;  // Print  $\log(\text{abs}(\text{Re}((\nu + m)A_m/z))$ 
    bool printLogImAd;  // Print  $\log(\text{abs}(\text{Im}((\nu + m)A_m/z))$ 
    bool printArgAd;    // To print  $\arg((\nu + m)A_m/z)$ 
    bool printReAd;     // To print  $\text{Re}((\nu + m)A_m/z)$ 
    bool printImAd;     // To print  $\text{Im}((\nu + m)A_m/z)$ 
    // Options to print full precision coefficients to stdout
    bool coutA;         // Option to print  $A_m$ 
    bool coutReA;       // Option to print  $\text{Re}(A_m)$ 
    bool coutImA;       // Option to print  $\text{Im}(A_m)$ 
    bool coutAd;        // Option to print  $(\nu + m)A_m/z$ 
    bool coutReAd;      // Option to print  $\text{Re}((\nu + m)A_m/z)$ 
    bool coutImAd;      // Option to print  $\text{Im}((\nu + m)A_m/z)$ 
};

```

Definition of the OdeResults structure

```

struct OdeResults {
    cl_N* fu;          // Pointer to pre-created array fu[2]
    int maxAExponent;  // Largest term in function series
    int maxAdExponent; // Largest term in derivative series
    int emmAtAMax;     //  $m$  where  $\text{abs}(A_m)$  is largest
    int emmAtAdMax;    //  $m$  where  $\text{abs}((\nu + m)A_m)$  is largest
    int lengthOfSum;   // Number of terms summed
    double timeUsed;   // Time used for evaluation
    double lgErrorF;   // Estimated error in function
    double lgErrorFd;  // Estimated error in derivative
    int returnStatus;  // -1:  $s = 0$ 
                    // -2:  $z = 0$ 
                    // -3:  $(\nu_P - \nu_M)$  is integer  $\neq 1$ 
                    // -4:  $\text{emm} \geq \text{emmTooLarge}$  was reached
                    // 1: All  $\text{abs}(A_m) < \text{Estimated error}$ 
                    // 2:  $\text{emm} = \text{emmMax}$  was reached
};

```